CLAIMS

What is claimed is:

1.   A computer implemented method of managing computer

5   resources comprising:

accessing an amount of computer resources allocated to a workload of a

computer system;

monitoring computer resource usage of said workload; and

determining a range of computer resources to make available for other

10   use provided said computer resource usage of said workload exceeds said

amount of computer resources allocated to said workload, wherein said

monitoring and said determining occur within a user space.

2.   The method of Claim 1 wherein said computer resources comprises

15   physical memory.

3.   The method of Claim 2 wherein said determining comprises

determining if a page of physical memory utilized by said workload has been

accessed by said workload within a predetermined period of time.

20

4.   The method of Claim 3 wherein said determining if a page of

physical memory utilized by said workload has been accessed by said workload

within a predetermined period of time comprises determining if said page has

been accessed by said workload since a previous determination of whether said page had been accessed by said workload.

5.   The method of Claim 1 wherein said computer resources comprises virtual memory.

6.   The method of Claim 1 wherein said computer resources comprises central processing unit time.

7.   The method of Claim 1 wherein said computer resources comprises input / output space.

8.   The method of Claim 1 wherein said computer resources comprises network bandwidth.

9.   A computer implemented method comprising:
    monitoring usage of a computing resource utilized by a workload wherein said monitoring is performed by a process within a user space;
    responsive to exceeding a limit on utilization of said computing resource, decreasing usage of said computing resource by said workload.

10.   The method of Claim 9 wherein said computing resource comprises physical memory.

11.   The method of Claim 10 wherein said decreasing usage of said computing resource comprises paging a portion of said physical memory assigned to said workload out of said physical memory.

5   12.   The method of Claim 11 wherein said portion of said physical memory comprises a least recently used portion of said physical memory assigned to said workload.

13.   The method of Claim 9 wherein said decreasing usage does not

10   halt operation of said workload.

14.   The method of Claim 9 wherein said decreasing usage is initiated by a process of said workload.

15   15.   The method of Claim 9 wherein said process that performs said monitoring is not an operating system kernel process.

16.   A computer implemented method comprising:

accessing a list of memory pages assigned to said workload;

20   responsive to a request from a first process of said workload for memory which exceeds a predetermined memory limit for said workload, selecting a plurality of memory pages from said list of memory pages; and

initiating a second process within a user space to page out said plurality of memory pages and wherein said selecting occurs within said user space.

17.   The method of Claim 16 wherein said accessing, selecting and

5   initiating are performed by said second process within said workload.

18.   The method of Claim 16 wherein said second process is not an operating system kernel process.

10   19.   The method of Claim 16 wherein said second process is loaded into a user space.

20.   The method of Claim 16 wherein said plurality of memory pages comprises memory pages that are least recently used.

15

21.   The method of Claim 20 wherein said plurality of memory pages comprises memory pages that are least recently used by said workload.

22.   The method of Claim 20 wherein said page out of said plurality of

20   least recently used memory pages reduces a number of memory pages assigned to said workload to below said memory limit.

23.   The method of Claim 20 wherein said plurality of least recently used memory pages comprises the minimum number of memory pages to reduce

said number of memory pages assigned to said workload below said memory limit.

24.   The method of Claim 16 wherein at least a portion of said workload continues to operate subsequent to said initiating.

25.   The method of Claim 16 wherein said initiating is not performed by an operating system kernel process.

26.   A computer-usable medium having computer-readable program code embodied therein for causing a computer system to perform a method, said method comprising:

accessing an amount of computer resources allocated to a workload of a computer system;

monitoring computer resource usage of said workload; and

determining a range of computer resources to make available for other use provided said computer resource usage of said workload exceeds said amount of computer resources allocated to said workload, wherein said monitoring and said determining occur within said workload.

27.   The computer-usable medium of Claim 26 wherein said computer resources comprise memory.

28.   The computer-usable medium of Claim 27 wherein said memory comprises physical memory.

29.   The computer-usable medium of Claim 28 wherein said
5   determining comprises determining if a page of physical memory utilized by said workload has been accessed by said workload within a predetermined period of time.

30.   The computer-usable medium of Claim 29 wherein said
10   determining if a page of physical memory utilized by said workload has been accessed by said workload within a predetermined period of time comprises determining if said page has been accessed by said workload since a previous determination of whether said page had been accessed by said workload.

15   31.   The computer-usable medium of Claim 30 wherein a page is identified as least recently used if said page has not been accessed by said workload since a previous determination of access for said page indicated that said page had been accessed by said workload.

20   32.   A computer implemented method comprising:
        scanning pages for a first workload of a computer to determine if each of said pages was accessed since a last scan;
        setting bits indicative of the result of said scanning within a scoreboard related to said first workload; and

repeating said scanning and said setting for a second workload, wherein said scanning and said setting occur within a user space of said computer.

5   33.   The method of Claim 32 wherein said scanning comprises checking said pages according to an order inherent to a list of said pages.

34.   The method of Claim 33 further comprising paging out a plurality of pages utilized by said first workload responsive to said determining.

10

35.   The method of Claim 32 further comprising determining if the number of pages utilized by said workload exceeds a predetermined limit.

36.   A computer implemented method comprising:

15   accessing memory usage for a workload and examining page usage for each process of said workload;

aggregating usage of said each process to determine an aggregate usage for said workload;

if said aggregate usage does not exceed a memory utilization limit

20   for said workload, repeating said accessing and aggregating for a next workload;

if said aggregate usage exceeds said memory utilization limit for said workload, determining least recently used pages based on accessed bits associated with said workload;

if said aggregate usage exceeds said memory utilization limit for said workload, supplying a range of least recently used pages in a system call to an operating system kernel for evicting said range of least recently used pages to reduce resource usage by said workload; and

5   retaining at least partial operation of said workload during said page evicting.

37.   The method of Claim 36 wherein said determining and said supplying occur in a plurality of user space processes.

10

38.   A computer implemented method of managing computer resources over a plurality of workloads, said method comprising:

for each workload of said plurality of workloads, monitoring respective workload resource usage against a respective allotment of each workload;

15   determining a range of computer resources to page out for each workload whose resource usage exceeds its respective allotment; and

initiating a paging out operation of said range of computer resources and wherein said monitoring, said determining and said initiating all occur within a process of user space.

20

39.   The method as described in Claim 38 wherein said determining comprises determining least recently used pages for each workload whose resource usage exceeds its respective allotment.

40. The method as described in Claim 38 wherein said process is situated within a workload of said plurality of workloads.

41. The method as described in Claim 38 further comprising paging out said range of computer resources and wherein each workload whose resource usage exceeds its respective allotment remains partially operable during said paging out of its respective range of computer resources.

42. A computer system comprising:

a bus for functionally coupling elements of said computer system;

physical memory coupled to said bus for storing processor instructions and data;

a processor coupled to said bus and for implementing a method, said method comprising:

accessing an amount of computer system resources allocated to a workload of said computer system;

monitoring computer resource usage of said workload; and

determining a range of computer resources to make available for other use provided said computer resource usage of said workload exceeds said amount of computer resources allocated to said workload, wherein said monitoring and said determining occur within said workload.

43. The computer system of Claim 42 wherein said computer resources comprises physical memory.

44. The computer system of Claim 42 wherein said computer resources comprises virtual memory.

45. The computer system of Claim 42 wherein said computer resources comprises central processing unit time.

46. The computer system of Claim 42 wherein said computer resources comprises input / output space.

47. The computer system of Claim 42 wherein said computer resources comprises network bandwidth.

48. A computer system comprising:

first and second computers coupled via a network;

said first computer for implementing a method, said method comprising:

accessing an amount of computer system resources allocated to a workload of said second computer;

monitoring computer system resource usage of said workload;

determining, at said first computer, a plurality of computer system resources to make available for other use provided said computer

resource usage of said workload exceeds said amount of computer resources allocated to said workload; and

making, at said second computer, at least one of said plurality of computer system resources available for other use.

5

49.   The computer system of Claim 48 wherein said computer resources comprises physical memory.

50.   The computer system of Claim 48 wherein said computer resources

10   comprises virtual memory.

51.   The computer system of Claim 48 wherein said computer resources comprises central processing unit time.

15   52.   The computer system of Claim 48 wherein said computer resources comprises input / output space.

53.   The computer system of Claim 48 wherein said computer resources comprises network bandwidth.

20